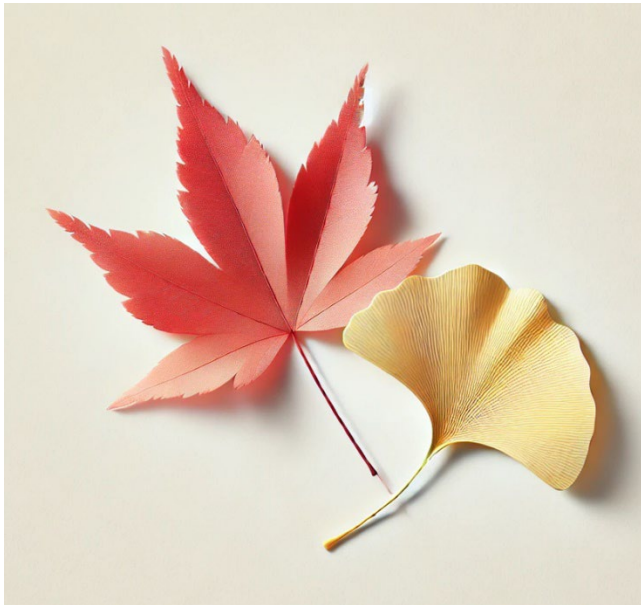


Assignment 4



BigData Week5

2025. 4.3

Eunhui Kim (김은희)

ehkim@kisti.re.kr

Assignment 4

- ❖ Submission due : April. 10 th, 23:55
- ❖ What to submit : Notebook file (.ipynb)
 - Colab : [File]-[Download]-[Download .ipynb]
 - Kaggle : [File]-[Download Notebook]

Notice

Please submit your code and answers to the reflection questions in a Jupyter notebook. Make sure to include explanations for each part of your code, including what you observe from the results.



Assignment 4

Problem 1. [Linear Regression With Diabets Dataset]

(slide p.41~p.43) As you can see the regression score, it is too low.

How can you get the higher regression score?

Make an higher regression score code referring to the given code.

And Explain more than one line why you choose the method and Explain the results by comparing original code and your own.



Assignment 4

Problem 2. [Naïve Bayes Classification With Meeting Meet-Up

(meeting_data_200.csv)]

(slide p.51~p.54) As you can see the classification accuracy for test dataset, it is too low.

How can you get the higher classification accuracy?

Make an higher naïve Bayesian classification accuracy code referring to the given code.

And Explain more than one line why you choose the method and Explain the results by comparing original code and your own.



Assignment 4

Problem 3. [Decision Tree with petrol Consumption Dataset]

(slide p.61~p.64) As you can see the metrics Mean Absolute Error/Mean Squared Error/Root Mean Squared Error, it is too high.

How can you get the lower error rate?

Make an lower error rate referring to the given code.

And Explain more than one line why you choose the method and

Explain the results by comparing original code and your own.

Hint: refer to slide 50.

You can tune the hyper-parameter of the model such as max_depth, max_features etc.

Reference :

[sklearn.tree.DecisionTreeRegressor — scikit-learn 1.3.2 documentation](https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeRegressor.html)



Q & A

김은희 (ehkim@kisti.re.kr)

TA: Yesim Selcuk (yesimselcuk@kisti.re.kr)